

CORNELL DATA SERVICES APRIL 2026 NEWSLETTER

COMPREHENSIVE DATA PLANNING & SERVICES

[Cornell Data Services](#) is a collaborative, campus-wide organization that assists with data management and sharing plans, applying best practices for managing data, and finding data-related services at any stage of the research process.

IN THIS NEWSLETTER:

[Spotlight: GitHub Repository vs. Research Data Repository](#)

[Upcoming webinars, training, and conferences](#)

[Funder and publisher news](#)

[Data in the news](#)

[In case you missed it](#)

SPOTLIGHT: GITHUB REPOSITORY VS. RESEARCH DATA REPOSITORY

Have you seen the term “repository” used in more than one way? If you work in GitHub, you’re likely familiar with repositories as spaces for code and collaboration. You may have also heard of [research data repositories](#) for long-term preservation and sharing.

This overlap in terminology can be confusing and sometimes leads users to think that these two types of repositories are the same, but they have quite different purposes.

In [GitHub](#), a “repository” is a space for collaborating with others and tracking code and file changes over time. While collaboration spaces such as GitHub may be sufficient for short-term sharing, they are not great choices for the long term. In contrast, research data repositories provide long-term access to and preservation of data, software, and code and support the [FAIR Guiding Principles](#) (Findable, Accessible, Interoperable, and Reusable). Depositing in research data repositories also helps researchers get a DOI and meet data-sharing requirements from federal funders and publishers.

The table below highlights key differences between the two.

| | GitHub Repository | Research Data Repository |
|------------------------|--|--|
| Primary purpose | Code development, collaboration, and version control | Long term access, sharing and preservation of research outputs |

| | | |
|--|---|---|
| Typical content | Source code, scripts, documentation, project files and small datasets | Final research data, software releases, publications, supplementary materials |
| Metadata and documentation | Minimal and basic (e.g., only a README file); does not follow FAIR principles | Rich, standardized metadata; supports README files; follows FAIR principles |
| Funder and publisher requirements | Typically does not meet funder or publisher requirements | Meets funder and publisher requirements |
| Access and sharing | Open by default, but controlled by user or organization of the repository | Often open and made broadly accessible |
| Citations | Offers a CITATION.cff file; does not provide a persistent identifier (e.g., DOI) | Often provides citation with submission including a persistent identifier (e.g., DOI) |
| Persistent identifiers | URLs only (might change over time) | Persistent identifiers (e.g., DOIs) |
| Versioning | Content can change frequently and iteratively (e.g., commits) | Content is typically fixed or versioned (new versions are added, not replaced) |
| Curation | Typically, no formal curation layer; depends on the user | Often includes some level of curation by a curation team or repository maintainers (e.g., standardized metadata, documentation) |
| Preservation actions | Repositories can be deleted; migration and integrity checks are dependent on the user | Integrity checks, open file format recommendations, persistent identifiers; is institutionally or organizationally supported |
| Best use in research workflows | Development and collaboration space | Archival record for reuse and citation |

There are many options for sharing research data and code, and CDS can help you select an appropriate repository to ensure code remains discoverable, accessible, and preserved as part of the scholarly record.

We recommend [archiving a GitHub repository using Zenodo](#) (and submitting to the [Cornell University Research Data Community](#)) to ensure permanency and get a citable DOI; depositing code or scripts in [Cornell's eCommons repository](#) is another good alternative.

Email data-help@cornell.edu to set up a consultation or to ask questions.

WEBINARS | TRAINING | CONFERENCES

DATA HAPPY HOUR

Wednesday, April 29 | 4:30-6:30 pm | Big Red Barn

For graduate students from any discipline: Representatives from Cornell Data Services and the Library will host Data Bingo with catered food, drinks and prizes followed by an informal data tabling session afterward. The event is sponsored by the local Ithaca IEEE chapter, but all disciplines are welcome. [Learn more.](#)

TIB KNOWLEDGE LOOM WORKSHOP

Thursday, April 23 | 3-5 pm | Mann Library's Stone Classroom Rm 103

A workshop using the TIB Knowledge Loom, an Open Science digital library of analysis-ready scientific knowledge. Of particular interest for researchers working in R or Python. Co-sponsored by CSCU, CDS, and the Library. [Learn more.](#)

FINAL SESSIONS: [ART + FEMINISM 2026](#)

Art + Feminism is a global Wikimedia Project, which aims to address information gaps on Wikipedia and Wikidata. This year, in addition to opportunities to edit Wikipedia, Cornell University Library is hosting a panel discussion about this year's theme: **We need feminist data to tell feminist stories.** Help close the gender gap on Wikipedia! Join us on the following dates for workshops and communal editing events focused on gender, feminism, and the arts.

[Drop-in Wikipedia Editing](#)

Tuesday, April 21 | 2:30-4:30pm | Mann Library 102

[Drop-in Wikipedia Editing](#)

Thursday, April 23 | 3:30-5:30pm | Olin Library 107

DIGITAL HUMANITIES INITIATIVE EVENTS IN SPRING 2026

[The Digital Humanities Initiative at Cornell](#) will host several public events featuring researchers engaging digital humanities work. Lunch will be provided at the events that take place at 12pm. All are welcome:

- **April 24** | 12-1:30 p.m. | Olin Library Room 703
Graduate student panel, featuring Stephanie Belina (Communication), Iris Luo (Human Centered Design) and Sanghoon Oh (Literatures in English)

THOUGHT SUMMITS ON DATA SCIENCE TO BUILD RESILIENCE AND IMPROVE HUMANITARIAN RESPONSE

May 11-13

Hosted by CIDER in partnership with CDSES, this summit, which includes two public lectures, will focus on innovative uses of ever-expanding sources of data, including satellite imagery, weather forecasts, censuses, surveys, community-generated data, and machine learning to improve early warning systems, target humanitarian assistance, and design more impactful aid policies. [Learn more.](#)

FUNDER AND PUBLISHER NEWS

NSF DMSP REQUIREMENT UPDATE

Reflecting changes announced in the [2026 PAPPG 24-1 Supplement 2](#), effective **April 27, 2026**, the Data Management and Sharing Plan format for NSF proposals will change from a PDF document to a webform within the Research.gov site. [Learn more](#).

If you have questions or would like someone to review your DMSP, please reach out to data-help@cornell.edu.

NIH DMSP REQUIREMENT UPDATE

NIH released [Updated Elements of an NIH Data Management and Sharing Plan](#), effective for grant applications with due dates on or after **May 25, 2026**. In this update, they shared a new draft [Data Management and Sharing Plan Format Page](#). The 2026 DMS format is significantly different than the 2023 DMS format. [Learn more](#).

If you have questions or would like someone to review your DMS plan, please reach out to data-help@cornell.edu.

CCSS GRANTS AND AWARDS

Rolling Deadline | [Apply for Roper/CCSS Data Conversion Grant](#) | Cornell faculty and graduate students can apply for up to \$10K in grant funding to convert datasets in obsolete file formats (e.g., column binary or ASCII) into modern formats.

May 15 Deadline | [Trevor Pinch QuIRI Innovation Award](#) | Recognizes a Cornell faculty member who demonstrates innovation in developing, using, or teaching qualitative methods. The recipient will give the Trevor Pinch Innovation Award talk in Fall 2026 and will also receive a \$500 honorarium.

DATA IN THE NEWS

FEATURED IN: ATOMIC DISTORTIONS REVEAL NEW CLUES ABOUT SUPERCONDUCTIVITY (CORNELL CHRONICLE)

“To understand why these structural changes are important for the material’s electronic properties, the researchers built a model using the data obtained by multislice electron ptychography to disentangle different types of coupled structural distortions such as tilting, stretching and bending in the nickel-oxygen bonds.” – [Cornell Chronicle](#)

- **Related dataset:** Bhatt, L., & Goodge, B. (2026). Datasets for "Structural modifications in strain-engineered bilayer nickelate thin films" [Data set]. Zenodo. <https://doi.org/10.5281/zenodo.18917988>
- **Related article:** Bhatt, L. et al. (2026). Structural modifications in strain-engineered bilayer nickelate thin films. Nature <https://doi.org/10.1038/s41586-026-10446-2>

IN CASE YOU MISSED IT

NEW! DATA AND CODE SHARING IN ZENODO, IMPROVED

CORNELL UNIVERSITY RESEARCH DATA COMMUNITY IN THE ZENODO REPOSITORY

<https://zenodo.org/communities/cornelldata>

Cornell researchers can now get free data curation within [CERN's Zenodo repository](#) via the [Cornell University Research Data Community](#). Depositing data and code in Zenodo may be a good option if you or your collaborators are already using it, you want code-sharing integration with GitHub, or for file sizes exceeding 10 Gb. You can submit data to the Cornell community as a draft or already published in Zenodo: [how to submit a published record to a community](#).

The Cornell University Zenodo Community is [managed and curated](#) by Cornell University data curators, members of the [Research Data and Open Scholarship team](#). If you have questions, contact us at open-scholarship@cornell.edu.

THOUGHT SUMMITS TO EXPLORE AI AND DATA SCIENCE FRONTIERS

Cornell University has selected three new [Thought Summits for 2026](#), advancing collaboration in artificial intelligence and data science across disciplines ranging from veterinary medicine and environmental justice to humanitarian response.

- **Data Science to Build Resilience and Improve Humanitarian Response**
May 11–15, 2026 | Location: Warren Hall 401
- **Community-Centered AI: Methodological Innovations for Environmental Justice, Labor Equity, and Data Sovereignty**
May 18–20, 2026 | Location: Computing and Information Science Building, Room 142
- **From Data to Animal Health: Building Benchmarks for AI-Driven Veterinary Innovation**
June 9–11, 2026 | Location: College of Veterinary Medicine

SCHEDULE A CONSULTATION

CDS [consultants](#) are available year-round to answer questions and help with data-related problems. We can meet via Zoom, email, or phone.

Send a request to data-help@cornell.edu to arrange a consultation at a time that works for you.

HAS THIS NEWSLETTER BEEN HELPFUL?

Tell us about it using this [feedback form](#) and we'll send you a **Cornell Data Services** goody bag!



 <https://data.research.cornell.edu>

 data-help@cornell.edu

Join or leave this list: <https://data.research.cornell.edu/contact-us/>